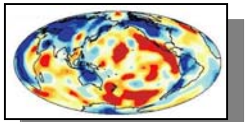
$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

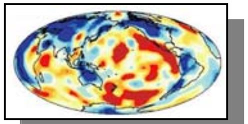
Inverse Problems in Geophysics



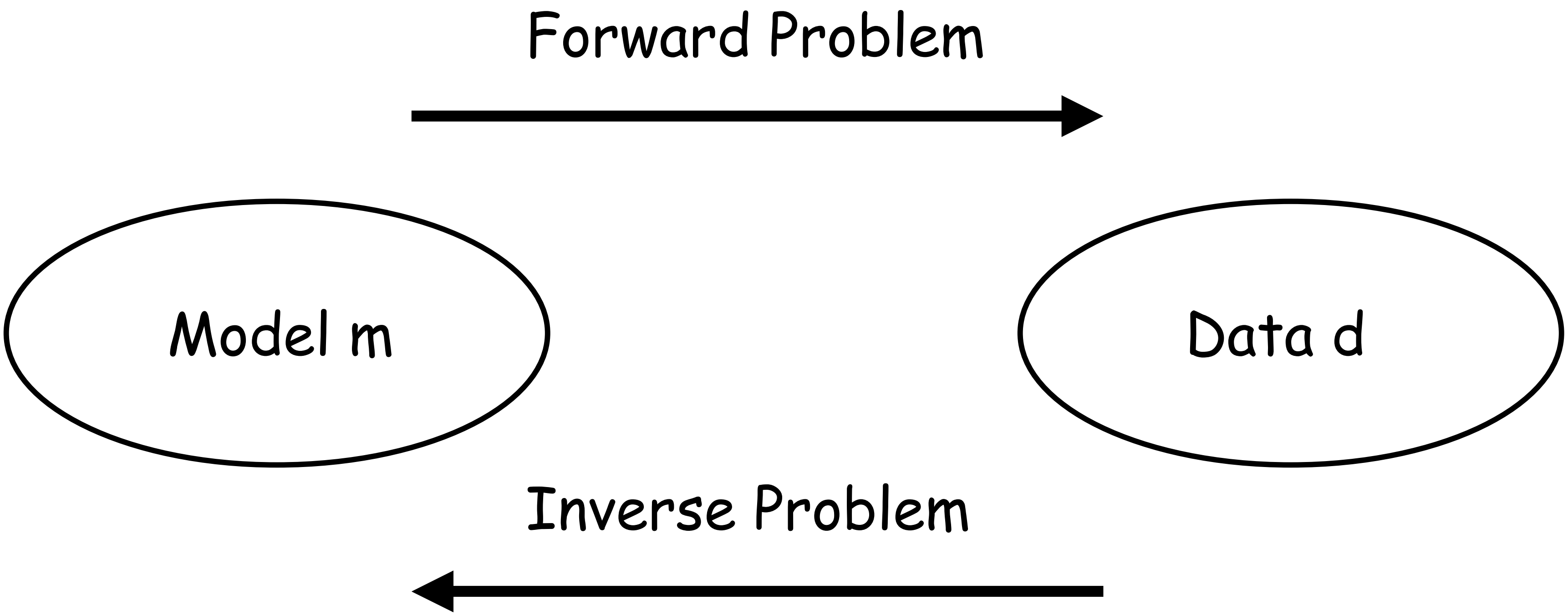
What is an inverse problem?
- Illustrative Example
- Exact inverse problems
- Linear(ized) inverse problems
- Nonlinear inverse problems

Examples in Geophysics
- Traveltime inverse problems
- Seismic Tomography
- Location of Earthquakes
- Reflection Seismology

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

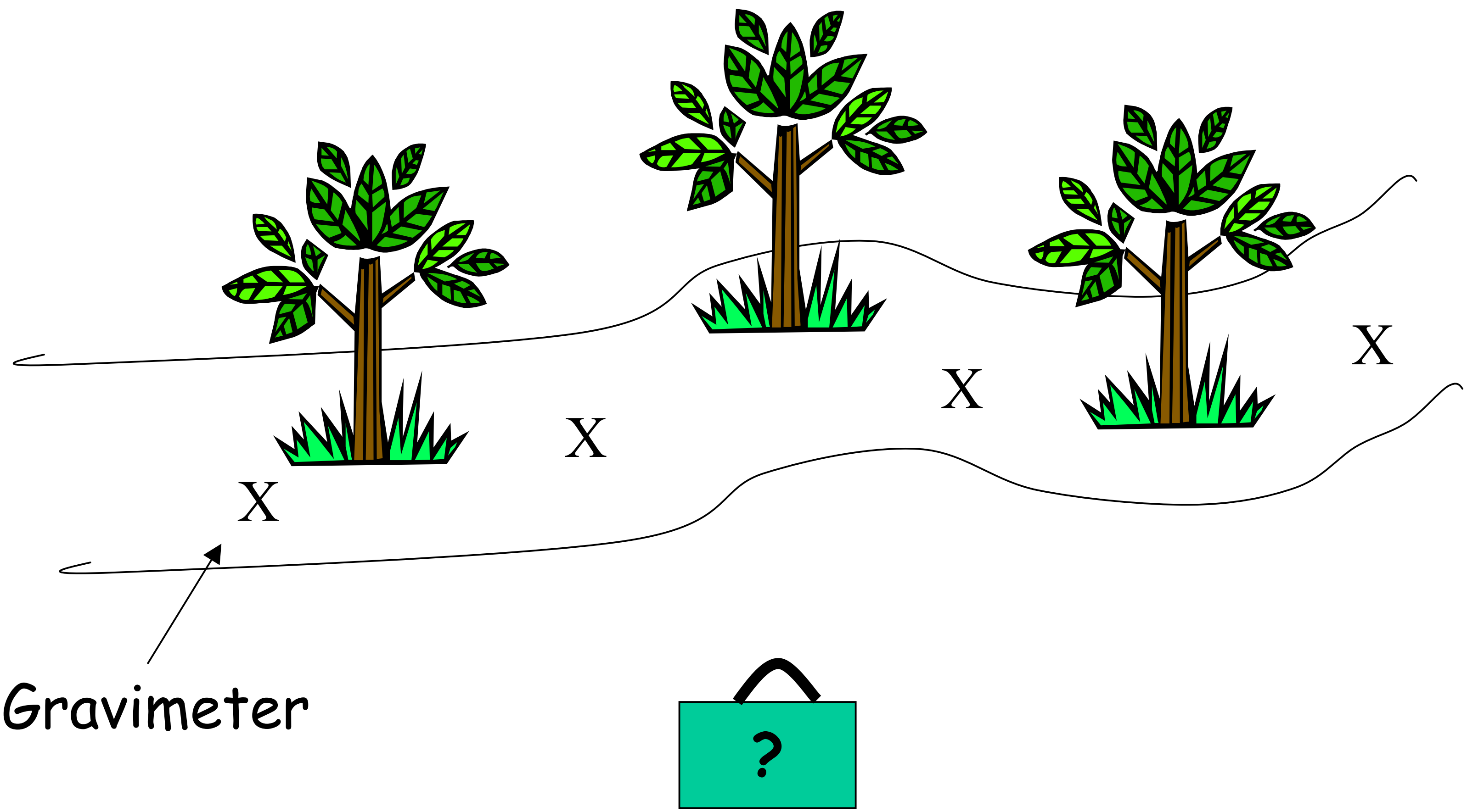


What is an inverse problem?

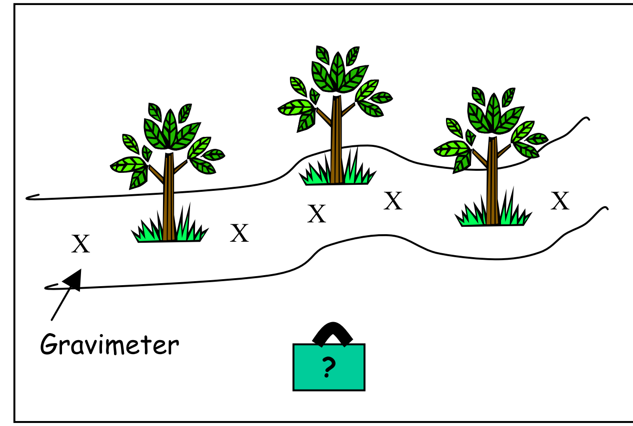


$$\sigma(d,m) = k\,\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Treasure Hunt



Treasure Hunt – Forward Problem



We have observed some values:

10, 23, 35, 45, 56 μgals

How can we relate the observed gravity values to the subsurface properties?

We know how to do the *forward* problem:

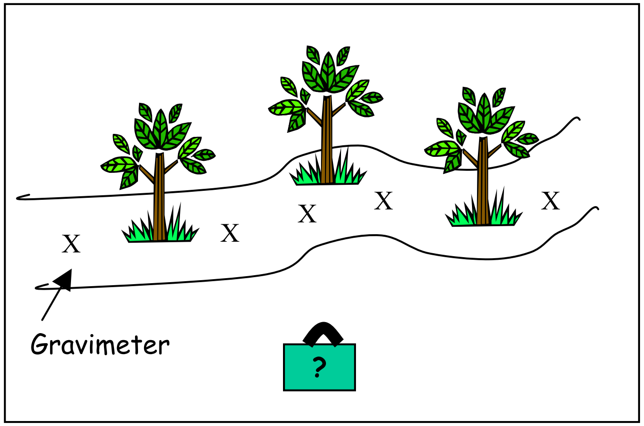
$$\Phi(r) = \int \frac{G\rho(r')}{|r - r'|} dV'$$

This equation relates the (observed) gravitational potential to the subsurface density.

-> given a density model we can predict the gravity field at the surface!

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Treasure Hunt – Trial and Error



What else do we know?

Density sand: 2,2 g/cm^3
Density gold: 19,3 g/cm^3

Do we know these values *exactly*?
How can we find out?
Where is the box with gold?

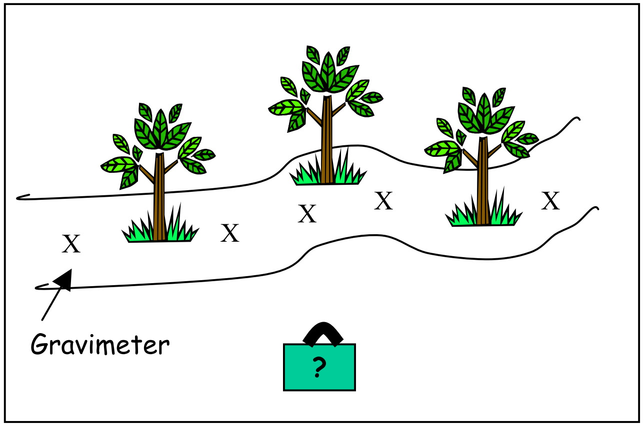
One approach:

Use the *forward* solution to calculate many models for a rectangular box situated somewhere in the ground and compare the *theoretical* (*synthetic*) data to the observations.

->Trial and error method

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Treasure Hunt – Model Space

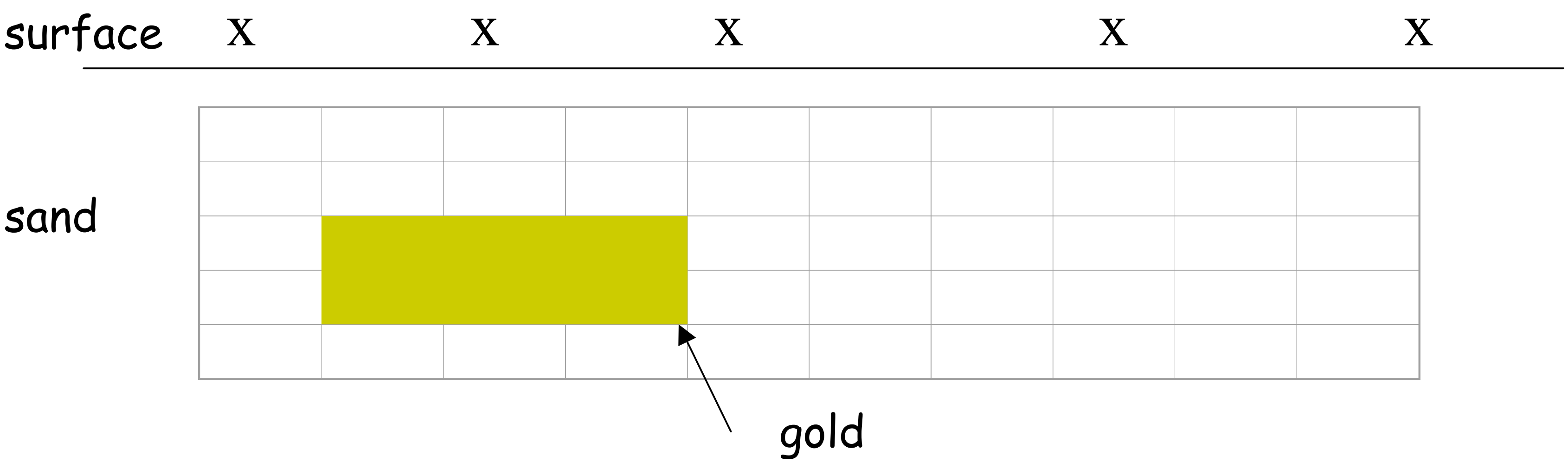


But ...

... we have to define *plausible* models for the beach. We have to somehow describe the model geometrically.

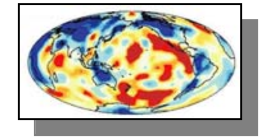
-> Let us

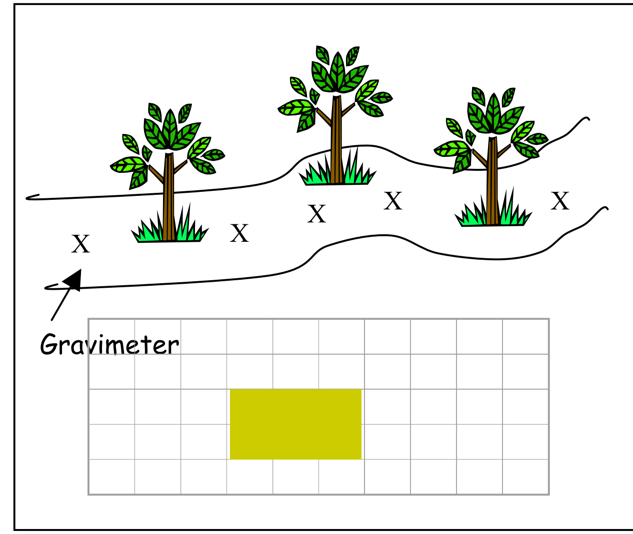
- divide the subsurface into rectangles with variable density
- Let us assume a flat surface



$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Treasure Hunt – Non-uniqueness





Could we go through all possible models and compare the synthetic data with the observations?

- at every rectangle two possibilities (sand or gold)
- $2^{50} \sim 10^{15}$ possible models

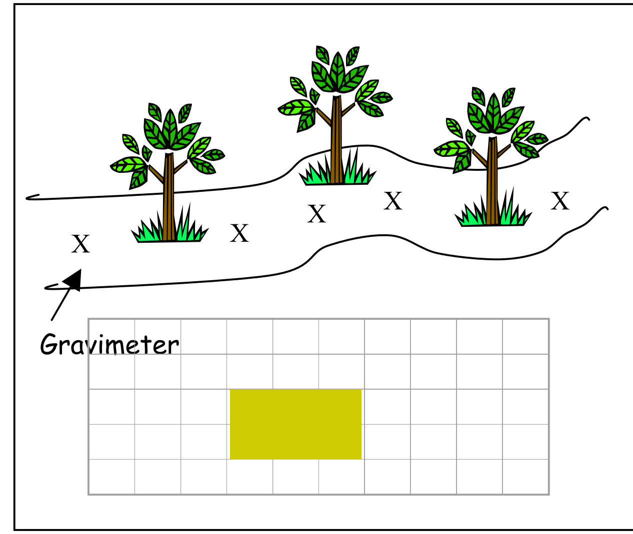
- **Too many models!**

- We have 10^{15} possible models but only 5 observations!
- It is likely that two or more models will fit the data (possibly perfectly well)

-> **Nonuniqueness of the problem!**

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Treasure Hunt – A priori information

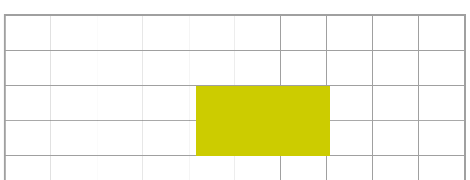


Is there anything we know about the treasure?

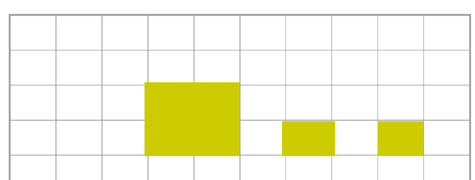
- How large is the box?
- Is it still intact?
- Has it possibly disintegrated?
- What was the shape of the box?
- Has someone already found it?

This is independent information that we may have which is as important and relevant as the observed data. This is called ***a priori*** (or prior) information. It will allow us to define plausible, possible, and unlikely models:

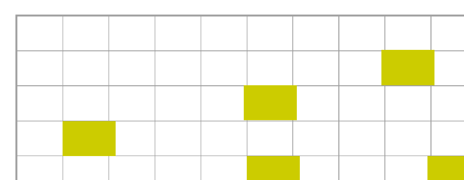
plausible



possible

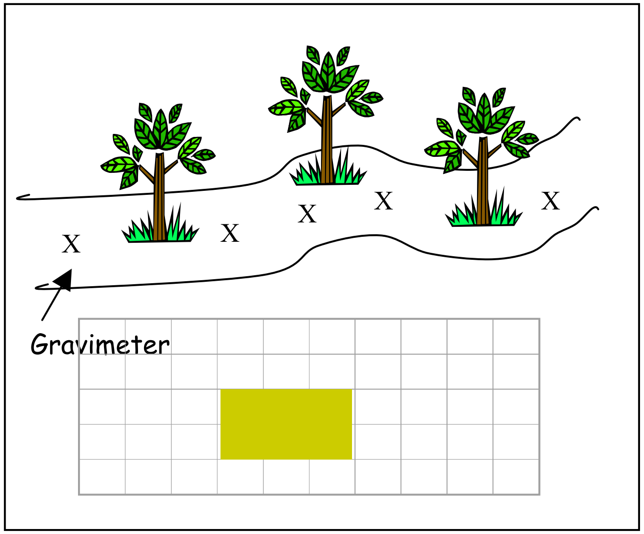


unlikely



$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Treasure Hunt – Uncertainties (Errors)



Do we have errors in the data?

- Did the instruments work correctly?
- Do have to correct for anything?
  (e.g. topography, tides, ...)

Are we using the right theory?

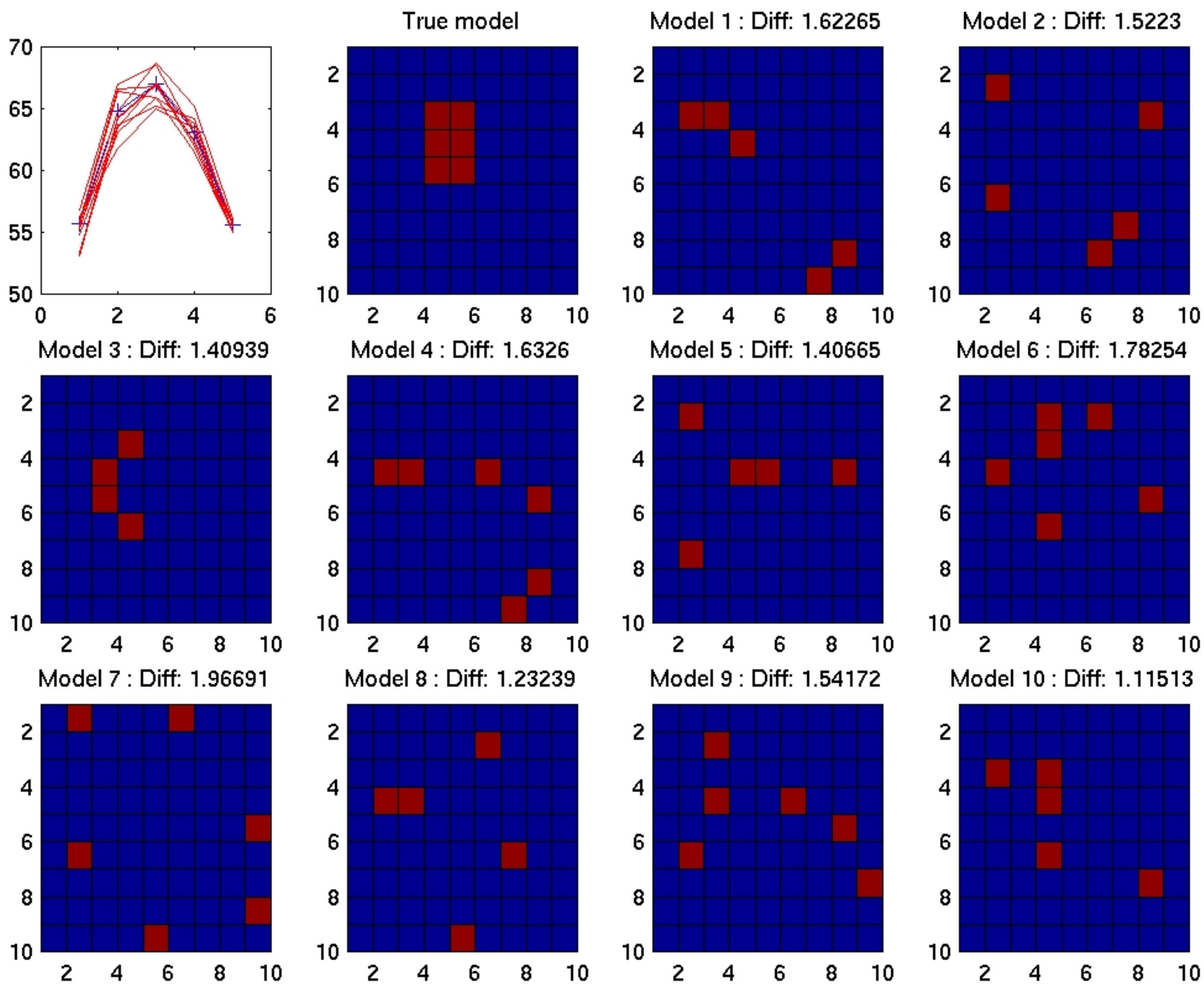
- Do we have to use 3-D models?
- Do we need to include the topography?
- Are there other materials in the ground apart from gold and sand?
- Are there adjacent masses which could influence the observations?

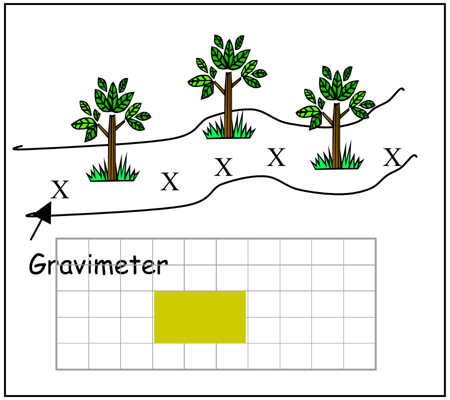
How (on Earth) can we quantify these problems?

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Treasure Hunt - Example

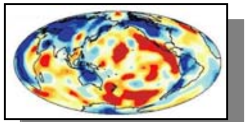
Models with less than 2% error.



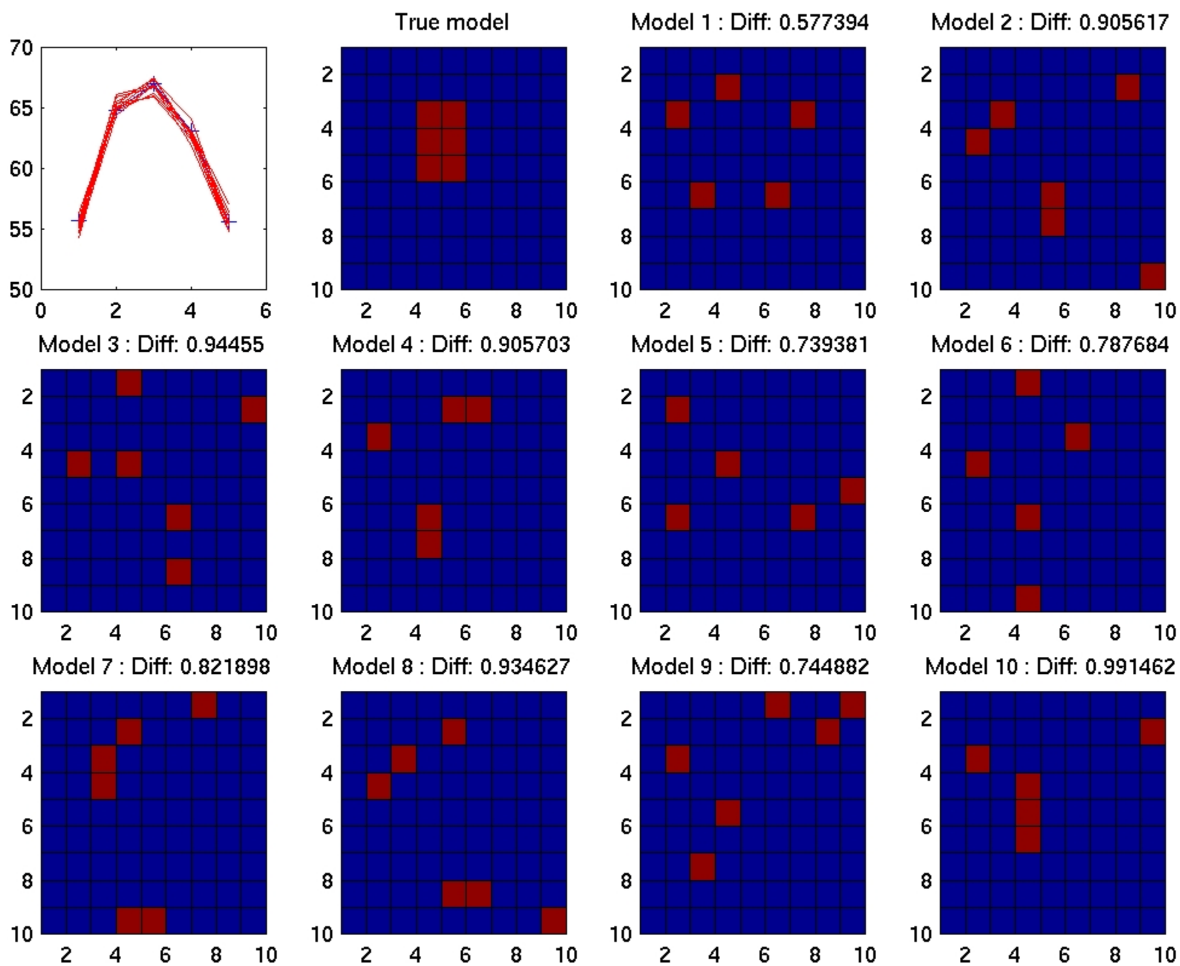


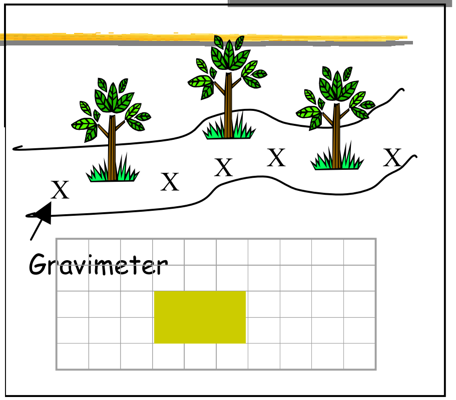
$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Treasure Hunt - Example



Models with less than 1% error.



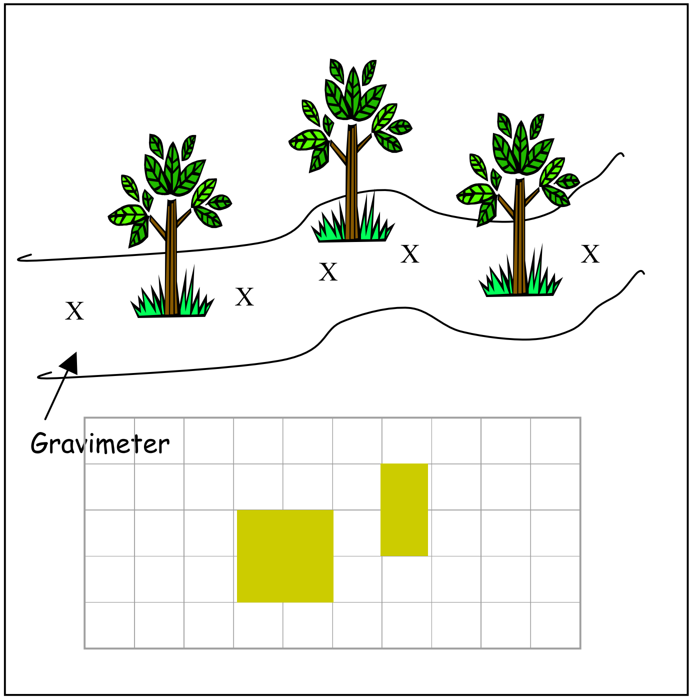


$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Treasure Hunt – Exercise

Exercise:

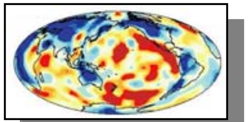
Now let us assume that we know the box has not disintegrated into less than two pieces. Change the calculations of the synthetic data and try to find the box, does it make a difference?

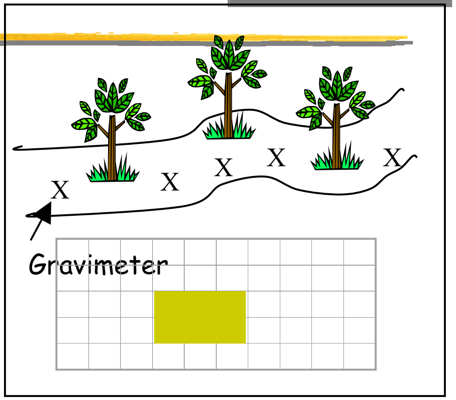


Parametrization of the box with two pieces

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Inverse Problems - Summary





Inverse problems – inference about physical systems from data

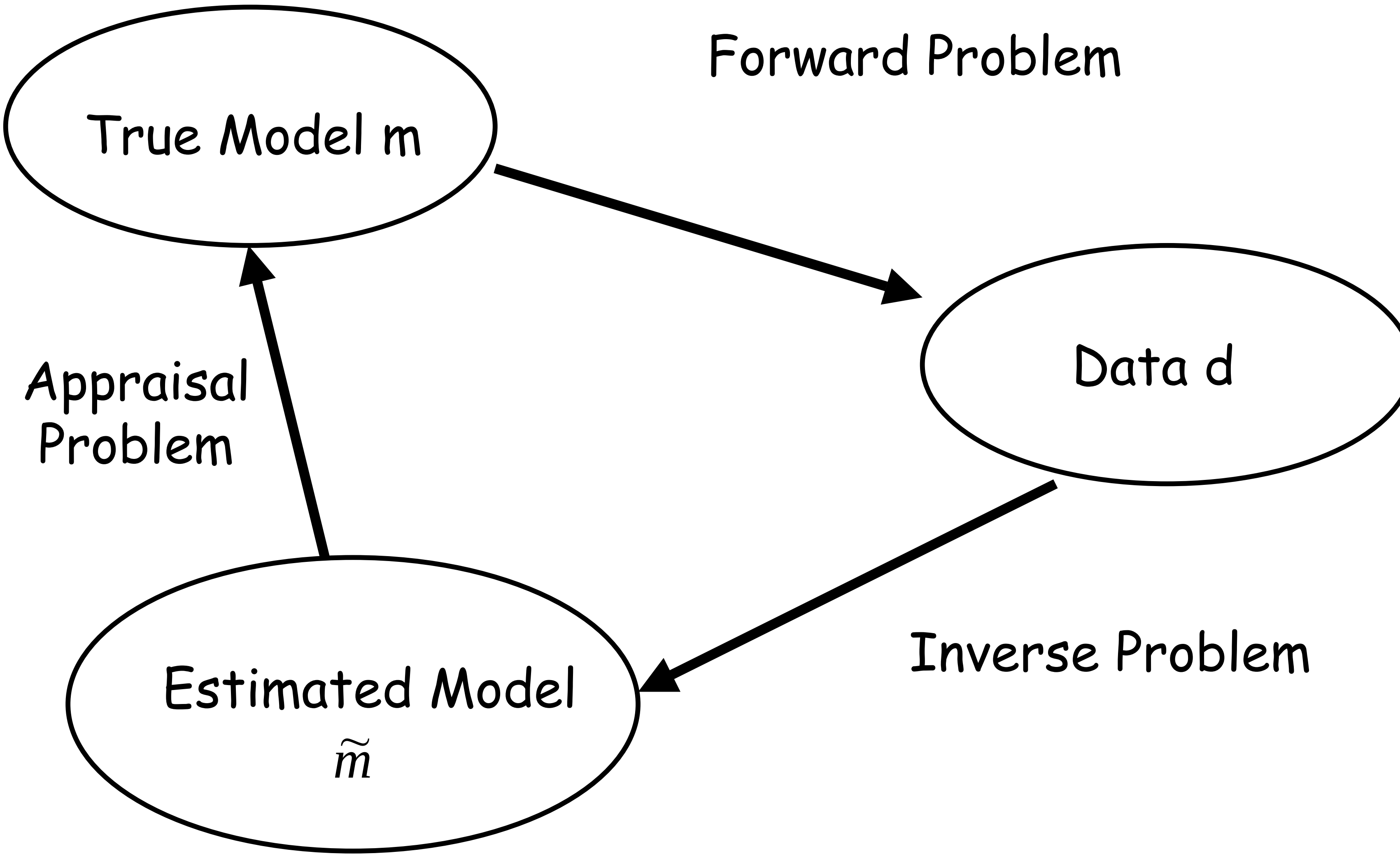
- Data usually contain errors (data uncertainties)
- Physical theories are continuous
- infinitely many models will fit the data (non-uniqueness)
- Our physical theory may be inaccurate (theoretical uncertainties)
- Our forward problem may be highly nonlinear
- We always have a finite amount of data

The fundamental questions are:

How accurate are our data?
How well can we solve the forward problem?
What independent information do we have on the model space (a priori information)?

$$\sigma(d,m) = k\,\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Corrected scheme for the real world

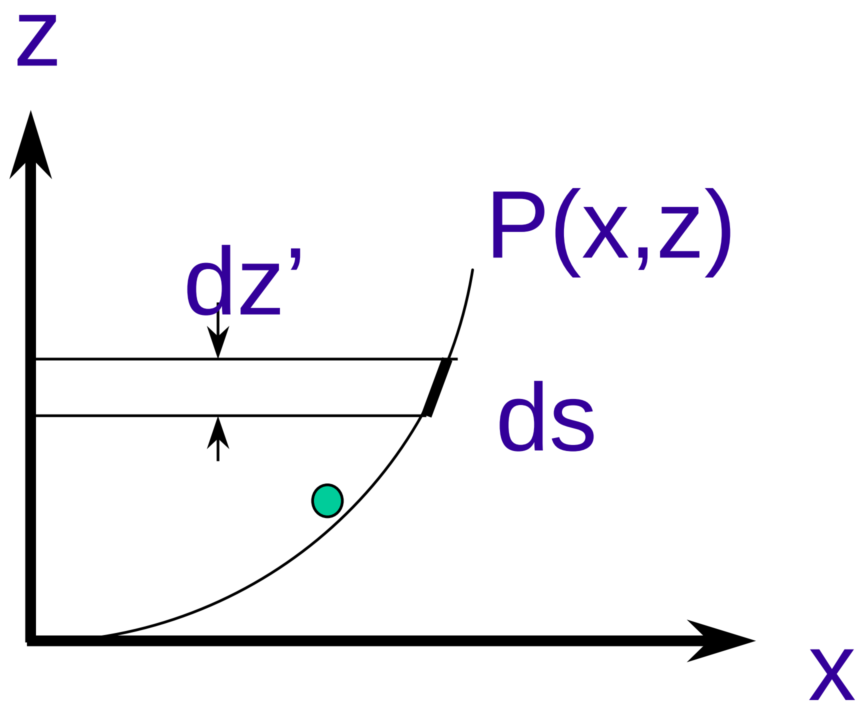


Exact Inverse Problems

Examples for exact inverse problems:

1. Mass density of a string, when all eigenfrequencies are known
2. Construction of spherically symmetric quantum mechanical potentials (no local minima)
3. Abel problem: find the shape of a hill from the time it takes for a ball to go up and down a hill for a given initial velocity.
4. Seismic velocity determination of layered media given ray traveltime information (no low-velocity layers).

Abel's Problem (1826)

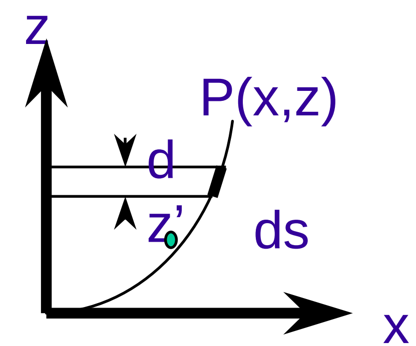


Find the shape of the hill !

For a given initial velocity and measured time of the ball to come back to the origin.

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

The Problem



At any point: $gz = \frac{1}{2}v_0^2$

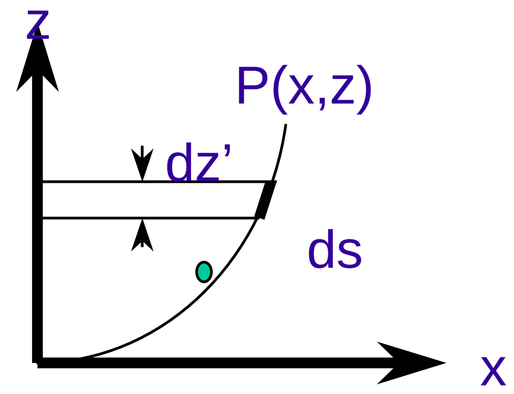
At z-z': $-mg(z-z') = \frac{1}{2}m(ds/dt)^2$

After integration: $t(z) = \int_0^z \frac{ds/dz'}{\sqrt{2g(z-z')}}dz'$

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

The solution of the Inverse Problem

$$\boxed{t(z)} = \int_0^z \frac{ds/dz'}{\sqrt{2g(z-z')}}dz'$$

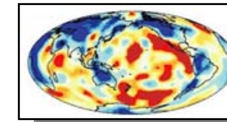


After change of variable and integration, and...

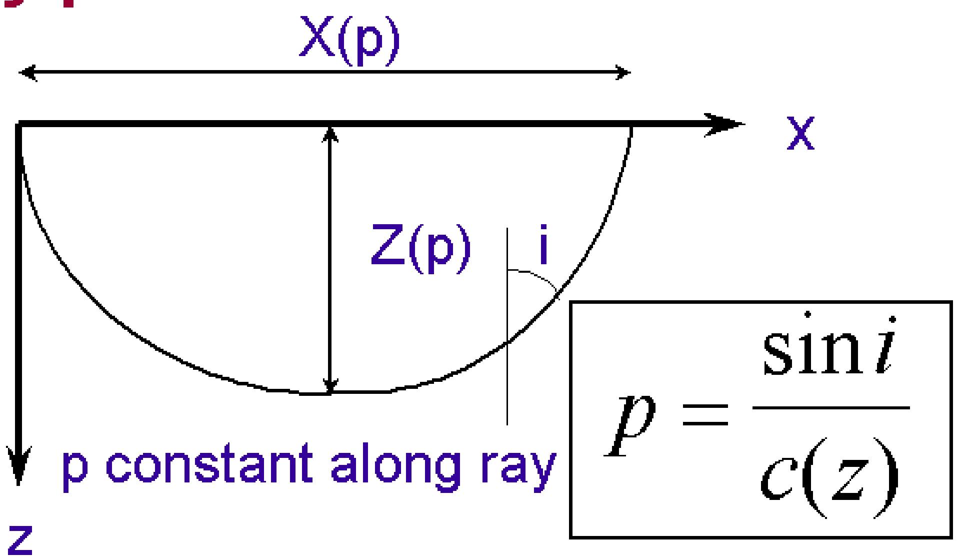
$$f(z') = -\frac{1}{\pi}\frac{d}{dz'}\int_{z'}^{a}\frac{\boxed{t(z)}\,dz}{\sqrt{z-z'}}$$

$$\sigma(d,m) = k\,\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

The seimological equivalent



Ray path in a flat Earth model



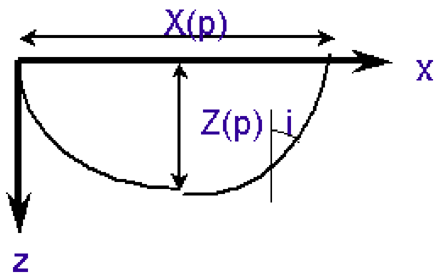
$$p = \frac{\sin i}{c(z)}$$

Find seismic velocities c(z) for observed X(p)

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Wiechert-Herglotz Method

Wiechert-Herglotz



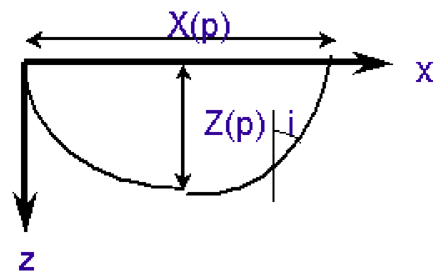
$$X(p) = 2\int_0^{Z(p)} \tan i \, dz$$

using p=sin*i*/c(z) ...

$$X(p) = 2\int_0^{Z(p)} \frac{p\,dz}{\sqrt{\boxed{c(z)}^{-2} - p^2}}$$

Distance and Travel Times

Wiechert-Herglotz



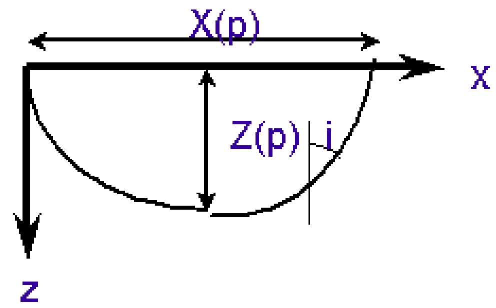
$$X(p) = 2\int_0^{Z(p)} \frac{p\,dz}{\sqrt{c(z)^{-2} - p^2}}$$

... or in terms of travel times

$$T(p) = 2\int_0^{Z(p)} \frac{c(z)^{-2}\,dz}{\sqrt{c(z)^{-2} - p^2}}$$

Solution to the Inverse Problem

Wiechert-Herglotz

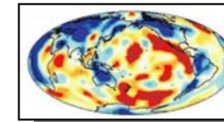


$$X(p) = 2\int_0^{Z(p)} \frac{p\,dz}{\sqrt{\boxed{c(z)}^{-2} - p^2}}$$

… after some manipulation …

$$\boxed{z(c)} = -\frac{1}{\pi}\int_{c_0^{-1}}^{c^{-1}} \frac{X(p)}{\sqrt{p^2 - c^{-2}}}\,dp$$

Wiechert-Herglotz Inversion



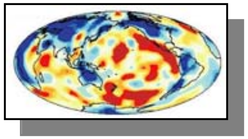
The solution to the **inverse** problem can be obtained after some manipulation of the integral :

$$T = p\Delta + 2\int_{r_0}^{r_1} \frac{\sqrt{r^2/c^2(z) - p^2}}{r^2}\,dr \Leftrightarrow \ln\left(\frac{r_0}{r_1}\right) = \frac{1}{\pi}\int_0^{\Delta_1} \cosh^{-1}\left(\frac{p}{\xi_1}\right) d\Delta$$

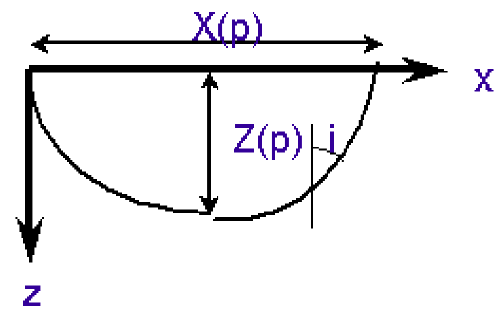
forward problem **inverse problem**

The integral of the inverse problem contains only terms which can be obtained from observed $T(\Delta)$ plots. The quantity $\xi_1=p_1=(dT/d\Delta)_1$ is the slope of $T(\Delta)$ at distance Δ_1. The integral is numerically evaluated with discrete values of $p(\Delta)$ for all Δ from 0 to Δ_1. We obtain a value for r_1 and the corresponding velocity at depth r_1 is obtained through $\xi_1=r_1/v_1$.

Conditions for Velocity Model



Wiechert-Herglotz

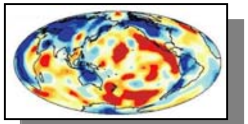


…conditions for X(p) and c(z)

- derivative of X(p) may be discontinuous
- X(p) must be continuous
- no low velocity channels
- rapid velocity increase is allowed

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Linear(ized) Inverse Problems



Let us try and formulate the inverse problem mathematically:
Our goal is to determine the parameters of a (discrete) model m_i, $i=1,\ldots,m$ from a set of observed data d_j $j=1,\ldots,n$. Model and data are functionally related (physical theory) such that

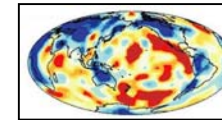
$$\begin{aligned}
d_1 &= A_1(m_1,\ldots,m_m)\\
d_2 &= A_2(m_1,\ldots,m_m)\\
&\vdots\\
d_n &= A_n(m_1,\ldots,m_m)
\end{aligned}$$

This is the nonlinear formulation.

Note that m_i need not be model parameters at particular points in space but they could also be expansion coefficients of orthogonal functions (e.g. Fourier coefficients, Chebyshev coefficients etc.).

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Linear(ized) Inverse Problems



If the functions $A_i(m_j)$ between model and data are linear we obtain

$$d_i = A_{ij} m_j$$

or

$$\boxed{\mathbf{d} = \mathbf{Am}}$$

in matrix form. If the functions $A_i(m_j)$ between model and data are mildly non-linear we can consider the behavior of the system around some known (e.g. initial) model m_j^0:

$$d_i = A_l(m_j^0) + \left.\frac{\partial A_i}{\partial m_j}\right|_{m_j^0} \Delta m_j + \ldots$$

Linear(ized) Inverse Problems

We will now make the following definitions:

$$d_i = A_i(m_j^0) + \Delta d_i$$

$$\Delta d_i = d_i - A_i(m_j^0)$$

$$\boxed{d_i = A_i(m_j^0) + \left.\frac{\partial A_i}{\partial m_j}\right|_{m_j^0} \Delta m_j + \ldots}$$

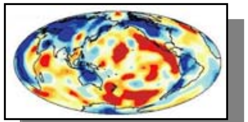
Then we can write a linear(ized) problem for the nonlinear forward problem around some (e.g. initial) model m_0 neglecting higher order terms:

$$\Delta d_i = \left.\frac{\partial A_i}{\partial m_j}\right|_{m_j^0} \Delta m_j \qquad \Delta d_i = A_{ij}\Delta m_j \qquad A_{ij} = \left.\frac{\partial A_i}{\partial m_j}\right|_{m_j^0}$$

$$\boxed{\Delta \mathbf{d} = \mathbf{A}\Delta \mathbf{m}}$$

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Linear(ized) Inverse Problems



$$\Delta \mathbf{d} = \mathbf{A}\Delta \mathbf{m}$$

Interpretation of this result:

1. m_0 may be an initial guess for our physical model
2. We may calculate (e.g. in a nonlinear way) the synthetic data $d=f(m_0)$.
3. We can now calculate the data misfit, $\Delta d=d-d_0$, where d_0 are the observed data.
4. Using some formal inverse operator A^{-1} we can calculate the corresponding model perturbation Δm. This is also called the gradient of the misfit function.
5. We can now calculate a new model $m=m_0+ \Delta m$ which will – by definition – is a better fit to the data. We can start the procedure again in an iterative way.

Nonlinear Inverse Problems

Assume we have a **wildly** nonlinear functional relationship between model and data

$$\mathbf{d} = g(\mathbf{m})$$

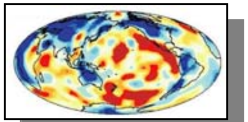
The only option we have here is to try and go – in a sensible way – through the whole model space and calculate the misfit function

$$\mathbf{L} = \left\| \mathbf{d} - g(\mathbf{m}) \right\|$$

and find the model(s) which have the minimal misfit.

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Model Search



The way how to explore a model space is a science itself!
Some key methods are:

1. Monte Carlo Method: Search in a random way through the model space and collect models with good fit.

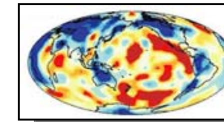
2. Simulated Annealing. In analogy to a heat bath, or the generation of crystal one optimizes the quality (improves the misfit) of an ensemble of models. Decreasing the temperature would be equivalent to reducing the misfit (energy).

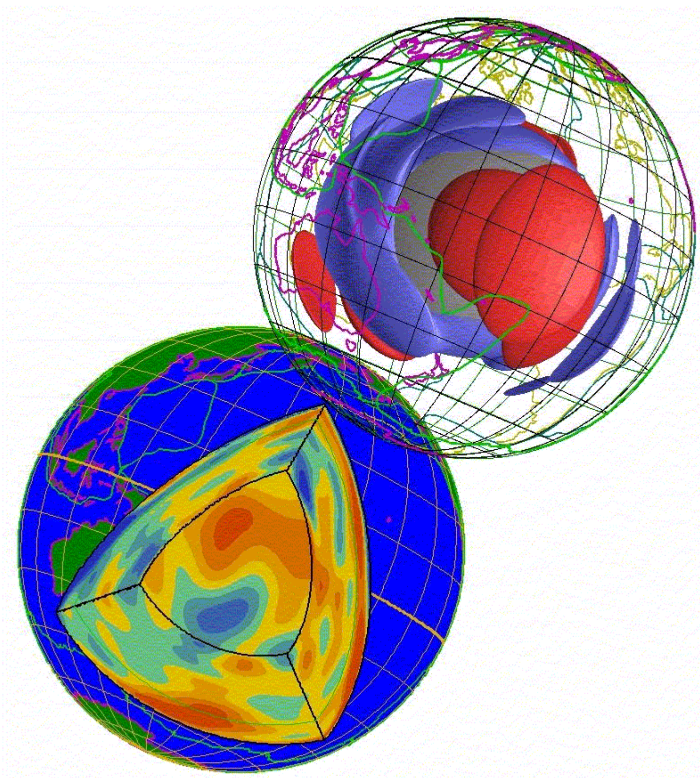
3. Genetic Algorithms. A pool of models recombines and combines information, every generation only the fittest survive and give on the successful properties.

4. Evolutionary Programming. A formal generalization of the ideas of genetic algorithms.

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Examples: Seismic Tomography





Data vector d:

Traveltimes of phases observed at stations of the world wide seismograph network

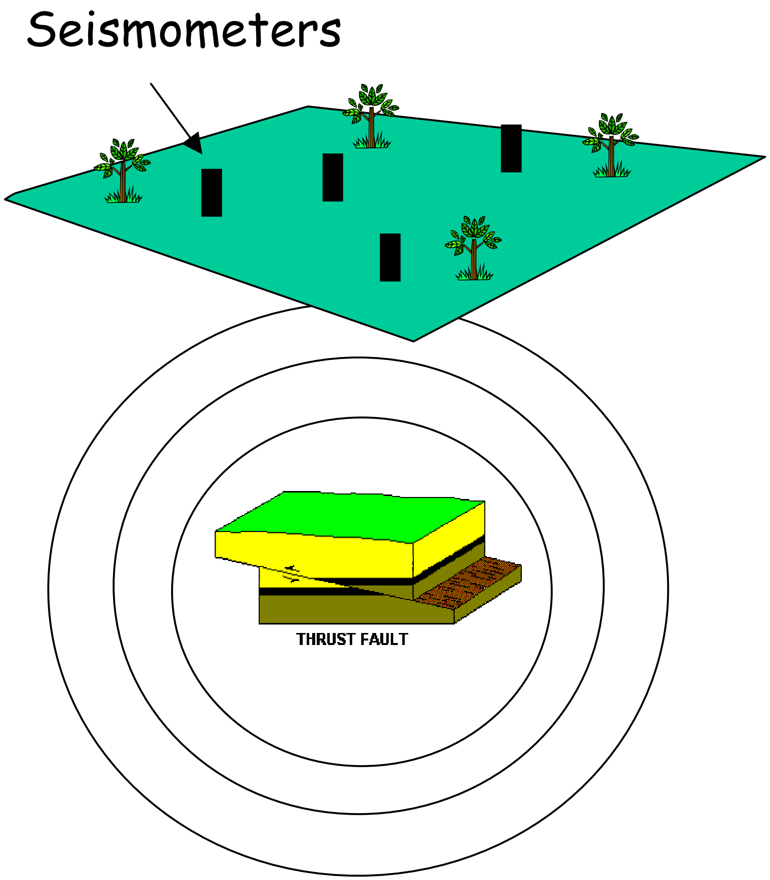
Model m:

3-D seismic velocity model in the Earth's mantle. Discretization using splines, spherical harmonics, Chebyshev polynomials or simply blocks.

Sometimes 10000s of travel times and a large number of model blocks: underdetermined system

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Examples: Earthquake location



Data vector d:

Traveltimes observed at various (at least 3) stations above the earthquake

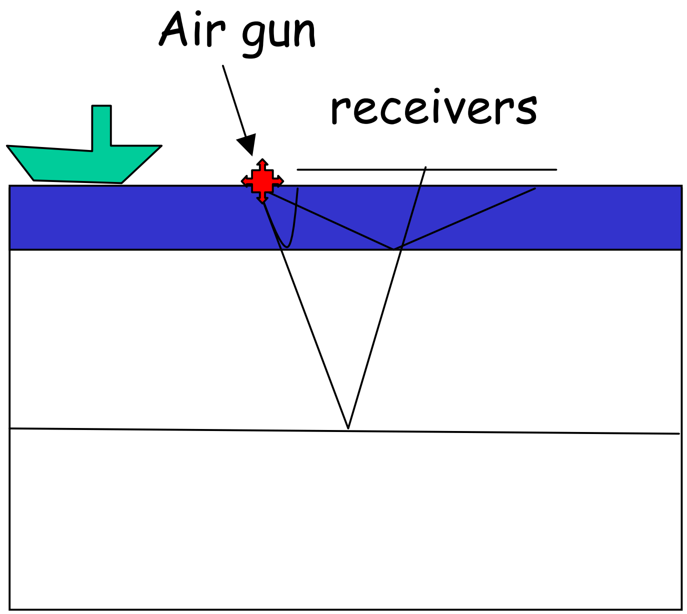
Model m:

3 coordinates of the earthquake location (x,y,z).

Usually much more data than unknowns: overdetermined system

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Examples: Reflection Seismology



Data vector d:

ns seismograms with nt samples

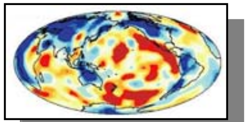
-> vector length ns*nt

Model m:

the seismic velocities of the subsurface, impedances, Poisson's ratio, density, reflection coefficients, etc.

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Inversion: Summary



We need to develop formal ways of

1. calculating an inverse operator for

   $$d = Am \;\; \text{->} \;\; m = A^{-1}d$$

   (linear or linearized problems)

2. describing errors in the data and theory (linear and nonlinear problems)

3. searching a huge model space for good models (nonlinear inverse problems)

4. describing the quality of good models with respect to the real world (appraisal).